The Predictive Power of the Yield Spread: Further Evidence and a Structural Interpretation[*]

Carlo A. Favero   Iryna Kaminska   Ulf Söderström

December 2004

Abstract

This paper brings together two strands of the empirical macro literature: the reduced-form evidence that the yield spread helps in forecasting output and the structural evidence on the difficulties of estimating the effect of monetary policy on output in an intertemporal Euler equation. We show that including a short-term interest rate and inflation in the forecasting equation improves the forecasting performance of the spread for future output but the coefficients on the short rate and inflation are difficult to interpret using a standard macroeconomic framework. A decomposition of the yield spread into an expectations-related component and a term premium allows a better understanding of the forecasting model. In fact, the best forecasting model for output is obtained by considering the term premium, the short-term interest rate and inflation as predictors. We provide a possible structural interpretation of these results by allowing for time-varying risk aversion, linearly related to our estimate of the term premium, in an intertemporal Euler equation for output.

Keywords: Yield curve, term structure of interest rates, predictability, forecasting, GDP growth, estimated Euler equation.

JEL Classification: E27, E37, E43.

---

1 Introduction

A large forecasting literature has examined variables that help predict the business cycle. The available empirical evidence tells us that the yield curve movements across the business cycle and yield curve fluctuations are a good leading economic indicator of GDP growth (Stock and Watson (1989), Harvey (1989), Estrella and Hardouvelis (1991), Plosser and Rouwenhorst (1994)). In particular, high spreads today are associated with higher GDP growth in the future. The more recent evidence tells us that yield spreads have become less useful as predictors in recent years (see, for example, Dotsey (1998)). In fact, one the spread's major predictive failures occurred on the occasion of the 1990–91 recession, incidentally immediately after the publication of some of the most influential articles cited above.

At the same time, there is also a vast literature discussing small-scale macroeconomic models in which aggregate demand is related to the real short-term interest rate. In models derived from a dynamic general equilibrium model in which output is determined by an intertemporal Euler equation (see, for example, Woodford (2003)), there is a relation between output and the expected path of future real short rates—the real long-term interest rate consistent with the expectations hypothesis—but there is no explicit role for the nominal yield spread.

Recently, such macro models consisting of a stylized demand and supply relationships closed by Taylor-type rules have been used to rationalize the declining predictive power of the yield spread for output growth (Feroli (2004), Estrella (2004)). By augmenting the standard three-equation setup with an expectations model of the term structure that makes the long-term interest rate equal to the average of expected future short rates, these authors show that the parameters in the forecasting relation between the yield spread and future economic growth depend on the form of the monetary policy reaction function. Hence, the declining coefficients in the relation between the yield spread and future economic growth can be related to the changing behaviour of the Fed, featuring more aggressive behaviour toward the gap between inflation and the target and less aggressive behaviour toward the output gap.

This rationalization is potentially interesting but it has one main shortcoming: the expectations theory of the term structure is assumed at the outset, so there is no role for any term premium. In an often quoted paper, Campbell and Shiller (1987) have shown that the yield spread can be decomposed into a weighted sum of future

1

expected changes in the short rate and a term premium. In principle, the spread can be large because future monetary policy is expected to be tight (as the central bank reacts to a higher expected level of activity in the future) or because the term premium is large as investors do not like to take on risk in bad times.

Hence, decomposing the yield spread into a term reflecting future monetary policy and a term reflecting the term premium could be useful to understand why yield curve fluctuations help in predicting subsequent economic activity and why the predictive power of the spread fluctuates over time. Moreover, such a decomposition would also allow to assess the importance of the term premium in determining macroeconomic fluctuations and hence whether there it might be useful to include the term premium in macroeconomic models.

However, such a decomposition is difficult in practice as it involves expectations about the future path of the short-term interest rate, and alternative decompositions may differ substantially depending on how expectations are modelled. Two recent studies decompose the yield spread to understand why it is a good predictor of real activity: Hamilton and Kim (2002) provide a decomposition using ex-post observed short rates to substitute for ex-ante expected rates, while Ang et al. (2004) use a VAR to project expectations for the short-term rate, the spread and GDP growth.

One shortcoming of these papers is that they do not mimic satisfactorily the process used in real time by market participants when forecasting short-term interest rates. The use of ex-post observed returns as a valid proxy for ex-ante returns has been questioned by Elton (1999), citing ample evidence against the belief that information surprises tend to cancel out over time. Hence, realized returns cannot be considered as an appropriate proxy for expected returns. In the procedure followed by Ang et al. (2004), the VAR is estimated on the full sample and projections are made in-sample. This procedure therefore cannot simulate the investors' effort to use the model in real time to forecast short rates, as the information from the whole sample is used to estimate parameters while investors can use only historically available information to generate predictions.

Our approach is to estimate a VAR at each point in time, using the historically available information, and then project short rates out-of-sample. Given the path of expected future short rates, we can construct yields to maturities consistent with the expectations theory and, as a residual, the term premium. Armed with the decomposition of the yield spread into expected future monetary policy and the term premium we aim at contributing first to the forecasting literature by evaluating the

2

role of our decomposition in understanding the predictive power of the spread for future GDP growth and, second, to the macroeconomic literature by providing an assessment of the importance of the direct inclusion of some measure of the term premium in models of aggregate demand.

The rest of this paper is organized as follows. Section 2 provides some stylized facts on the role of the spread, the short-term interest rate, the long-term interest rate and inflation in predicting future output fluctuations. Section 3 derives a decomposition of the yield spread into expected short rates and a measure of the term premium by following our real time procedure and compares our results with the measures derived by Hamilton and Kim (2002) and Ang et al. (2004). We then show how the different decompositions affect the predictive model for output. Section 4 provides some structural interpretation of our results, while Section 5 concludes.

2   Predicting GDP growth using interest rates

Throughout, we use quarterly U.S. data from 1954:1 to 2004:2 on the five-year Treasury bond yield, i_t^{20}; the three-month T-bill rate, i_t^1; real GDP, Y_t; and inflation measured by the GDP deflator, π_t. All data have been taken from the FRED database at the Federal Reserve Bank of St. Louis. We begin by analyzing predictive regressions for one-year ahead annual GDP growth ($\Delta_4 y_{t+4} = \log(Y_{t+4}) - \log(Y_t)$) based on possible combinations of a set of regressors containing the yield spread ($S_t^{20} = i_t^{20} - i_t^1$), the five-year rate, the three-month rate, and annual (four-quarter) inflation, for a total of 15 models. (All regressions also include a constant term.) In practice, we estimate all possible specifications of the forecasting equation

$$\Delta_4 y_{t+4} = \boldsymbol{\beta}_i' \mathbf{X}_{t,i} + \varepsilon_{t+4,i}, \tag{1}$$

where $\mathbf{X}_{t,i}$ is the set of regressors, observable at time t, included in the ith specification ($i = 1, \ldots, 15$) for future GDP growth. Our set of predictive models contains as special cases the standard predictive model for GDP growth based on the spread only and a model similar to a simple specification for the aggregate demand, relating GDP growth to the real ex-post short-term interest rate.

After twenty years of initialization we estimate all models recursively.[1] A range

---

[1]Thus the first regression uses information from the sample 1954:1–1973:4 to forecast GDP growth between 1973:4 and 1974:4, and the last regression uses the sample 1983:03–2003:02 to

of statistical selection criteria weighting goodness of fit against parsimony of the specification selects unequivocally the model based on the spread, the nominal short-term interest rate and inflation as the best predictive model. Figure 1 reports the recursively computed adjusted R^2 associated to five different predictive models of particular interest. Model 1 is the standard predictive model, including only the spread among the regressors; Model 2 includes the spread and the short-term rate; Model 3 includes the spread, the short-term rate and inflation; Model 4 includes the short-term rate and inflation; and Model 5 includes the long-term rate and inflation.

Model 3 uniformly dominates: while including the spread leads to an improvement of the forecasting performance for GDP growth (compare Models 3 and 4), the inclusion of the short-term nominal rate and inflation in addition to the spread causes a significant improvement in the forecasting performance (compare Models 1, 2 and 3).[2] Importantly, among these five models, the model most closely related to an aggregate demand equation (Model 4) dominates only a specification in which the long-term rate is substituted for the short-term rate (Model 5). Finally, we note that the adjusted R^2 displays a common downward trend for all models, consistent with the results in Dotsey (1998).

The indication given by the adjusted R^2 is confirmed by the pattern of the forecasting errors generated by the different models, reported in Figure 2. All models are unable to predict the 1990–91 recession, but models including the spread (Models 1–3) clearly dominate in the latter part of the sample. In particular, the specification which reflects the traditional aggregate demand equation (Model 4) features consistent predictive failures from the 1980s onward and clearly underperforms relative to the alternative models.[3]

The fact that the traditional specification of the aggregate demand equation (Model 4) underperforms so drastically relative to the models based on the spread naturally raises the question why the spread predicts GDP growth and how serious is the misspecification of structural models that find no role for this variable. To this end it is interesting to assess the economic significance of the coefficients in the model selected unequivocally as the best by the statistical selection criteria.

---

forecast GDP growth between 2003:02 and 2004:02.

[2]Ang et al. (2004) also note that including the short rate in addition to the spread increases the predictive power of the model.

[3]Similar evidence has been recently been labeled by Goodhart and Hofmann (2004) as "The IS puzzle."

4

Figure 3 reports recursive coefficient estimates with associated two-standard error confidence intervals for Model 3, that includes the spread, the short-term rate and annual inflation. These estimates show the well-known declining coefficient on the spread, a negative significant coefficient on the nominal short-term interest rate and a rarely significant but negative coefficient on inflation.

However, it seems difficult to make economic sense of this statistical model for future GDP growth: compared with standard aggregate demand equations that include the real short-term interest rate, the coefficient on inflation has the wrong sign, and the important role of the nominal yield spread is puzzling. To clarify this issue, we decompose the yield spread into two components: the expected future path of monetary policy relative to the current short rate and a term premium. Almost all rationalizations of the forecasting power of the spread have concentrated on the first expectations-related term, neglecting the role of the term premium, but in principle, fluctuations in term premia may be a powerful predictor of macroeconomic fluctuations. To explore more closely this issue we need a decomposition of the spread into the expectations-related and term premium components. Such decompositions are discussed and implemented in the next section.

3   Decomposing the yield spread

To decompose the yield spread into an expectations-related (ER) component and a term premium (TP), consider the following definition of the time-varying premium:

$$i_t^{20} = \frac{1}{20} \sum_{j=0}^{19} \mathrm{E}\left[i_{t+j}^1 \mid I_t\right] + TP_t, \tag{2}$$

where i_t^{20} is the five-year interest rate, i_t^1 is the three-month interest rate, I_t is the information available to agents when forming expectations at time t, and TP_t could be viewed as the sum of a liquidity premium and a risk premium. Following Campbell and Shiller (1987), equation (2) can be written in terms of the yield spread,

$S_t^{20} = i_t^{20} - i_t^1$ as

$$
\begin{aligned}
S_t^{20} &= \left\{ \frac{1}{20} \sum_{j=0}^{19} \mathrm{E}\left[i_{t+j}^1 \mid I_t \right] - i_t^1 \right\} + TP_t \\
&= \sum_{j=1}^{19} \frac{20-j}{j} \mathrm{E}\left[\Delta i_{t+j}^1 \mid I_t \right] + TP_t \qquad (3) \\
&= ER_t + TP_t.
\end{aligned}
$$

To decompose the yield spread, therefore, we first need a measure of the expected path of future short rates to construct the ER component. Then we can calculate the TP component as a residual term.

Recently, two papers—Hamilton and Kim (2002) and Ang et al. (2004)—have tried to assess the relative importance of the ER_t and TP_t terms in predicting future output growth. Having constructed measures of ER and TP, these authors estimate the regression

$$
\Delta_k y_{t+k} = \alpha_0 + \alpha_1 ER_t + \alpha_2 TP_t + \varepsilon_t. \qquad (4)
$$

However, as they use different techniques to decompose the yield spread, they obtain different results: Hamilton and Kim (2002) find that both the ER and TP components are significant in predicting output growth, while Ang et al. (2004) find that only the ER component is significant.

To decompose the yield spread, Hamilton and Kim (2002) construct a measure of the ER component using ex-post observed short rates instead of ex-ante expected rates. Thus, they write equation (4) as

$$
\Delta_k y_{t+k} = \alpha_0 + \alpha_1 \left\{ \frac{1}{20} \sum_{j=0}^{19} i_{t+j}^1 - i_t^1 \right\} + \alpha_2 \left\{ i_t^{20} - \frac{1}{20} \sum_{j=0}^{19} i_{t+j}^1 \right\} + u_t, \qquad (5)
$$

where the error term now is

$$
u_t = \varepsilon_t + (\alpha_2 - \alpha_1) \left\{ i_t^{20} - \frac{1}{20} \sum_{j=0}^{19} \mathrm{E}\left[i_{t+j}^1 \mid I_t \right] \right\}. \qquad (6)
$$

Equation (5) is then estimated by instrumental variables, exploiting the fact that under rational expectations the error term u_t should be uncorrelated with any variable known at time t. In particular, they estimate a just-identified model using as

6

instruments a constant, i_t^1, and i_t^{20}. As a result the measure of the TP component used by Hamilton and Kim (2002) is the fitted values from the first stage regression of $(i_t^{20} - \sum_{j=0}^{19} i_{t+j}^1 / 20)$ on a constant, i_t^1, and i_t^{20}.

Importantly, the Hamilton and Kim (2002) decomposition is constructed in a way that gives agents more information than they have when forecasting future short rates in real time: ER is constructed by first using perfect foresight, and then introducing expectational errors by running an IV procedure. Moreover, the relation between the instruments and the endogenous variables is estimated only once on the full sample. Therefore, at any earlier point in the sample the procedure exploits information not available to the agents at that time.

Ang et al. (2004) instead derive expectations for future short rates using a vector of state variables that follows a Gaussian Vector Autoregression with one lag:

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t. \tag{7}$$

The vector X_t contains two factors from the yield curve: the short rate i_t^1, expressed at a quarterly frequency, to proxy for the level of the yield curve, and the five-year spread, $i_t^{20} - i_t^1$, to proxy for the slope of the yield curve; and a macroeconomic factor: the quarterly rate of real GDP growth, Δy_t. Having estimated the VAR on the full sample, the expected short rate is calculated by simulating the VAR forward. Thus, as in Hamilton and Kim (2002), expectations are derived giving to agents some information (the full-sample coefficient estimates) that they cannot have in real time. This problem becomes particularly relevant when the parameters in the VAR are subject to shifts and structural breaks.

We therefore propose measures for ER_t and TP_t where we try to mimic the real-time forecasting procedure used by private agents. To construct these measures we estimate at each point in time, using the historically available information, the following model:[4]

$$X_t \;\; = \;\; \mu + \Phi(L)X_{t-1} + \Sigma \varepsilon_t, \tag{8}$$

---

[4]The adopted specification replicates closely that of Ang et al. (2004). We have also experimented an alternative specification which includes the first difference rather than the level of the short rate, a model that is consistent with the cointegrated VAR of Campbell and Shiller (1987). The results are rather robust, however the simulated values from the cointegrated model are more volatile than those in the baseline specification. We take this as evidence that the coefficients in the cointegrating vector may be different for some sample split from the unit values imposed in the specification with the spread.

where

$$X'_t = \left[i_t^1, i_t^{20} - i_t^1, \Delta_4 y_t, \pi_t \right].\tag{9}$$

We then simulate the estimated model forward, to obtain projections for all the relevant short rates and we construct ER as

$$\widehat{ER}_t = \sum_{j=1}^{19} \frac{20 - j}{20} \mathrm{E}\left[\Delta i_{t+j}^1 \mid \Omega_t \right],\tag{10}$$

where $\mathrm{E}\left[\Delta i_{t+j}^1 \mid \Omega_t \right]$ is the VAR-based projections for the future changes in the short rate, so Ω_t is the information set used by the econometrician to predict on the basis of the estimated VAR model. Importantly, in implementing our procedure the econometrician uses (almost) the same information available to market participants in real-time.[5] The expected future short rates at time t are constructed using information available in real time both for estimating the parameters and for projecting the model forward.

By combining equations (3) and (10) we obtain

$$
\begin{aligned}
S_t^{20} &= \left\{ \frac{1}{20} \sum_{j=0}^{19} \mathrm{E}\left[i_t^1 \mid I_t \right] - i_t^1 \right\} + TP_t \\
&= \sum_{j=1}^{19} \frac{20 - j}{20} \mathrm{E}\left[\Delta i_{t+j}^1 \mid I_t \right] + TP_t \\
&= ER_t + TP_t \\
&= \widehat{ER}_t + TP_t + \left(ER_t - \widehat{ER}_t \right),
\end{aligned}\tag{11}
$$

where

$$\widehat{ER}_t = \sum_{j=1}^{19} \frac{20 - j}{20} \mathrm{E}\left[\Delta i_{t+j}^1 \mid \Omega_t \right].\tag{12}$$

Equation (12) makes clear that deviations of S_t^{20} from $\widehat{ER}_t$ can be explained by movements in the term premium or by differences between the model-based forecasts, derived using the econometrician's information set Ω_t, and agents' expectations,

---

[5]Note that we use final data for GDP and the GDP deflator (not real-time data) when using the VAR to make real-time projections. This is in order to compare the results from the real-time decomposition with those from the HK and APW decompositions.

which are formed given the information set I_t, unknown to the econometrician. Under the assumption that this second term is negligible, significant deviations of S_t^{20} from $\widehat{ER}_t$ offer a measurable counterpart of the term premium.

Given that the statistical selection criteria favor Model 3 as the best for predicting output growth, we extend this model using the decomposed spread in the following way:[6]

$$\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 ER_t + \alpha_2 TP_t + \alpha_3 i_t^1 + \alpha_4 \pi_t + u_t. \tag{13}$$

We will consider three alternative measures of ER_t and TP_t. The first one is obtained by applying to our case the method proposed by Hamilton and Kim (2002), that is, by constructing the relevant variable ex-post and then by instrumenting it using the long-term rate and current GDP growth as additional instruments. Thus, in the first stage regressions ER_t and TP_t are regressed on a constant, i_t^1, π_t, i_t^{20}, and $\Delta_4 y_t$. We label these measures ER_t^{HK} and TP_t^{HK}. Note that when applying this method we need the last five years of observations to construct the ex-post measure, so given a total available sample of 1954:1–2004:2 we can derive the relevant measures for the sample 1954:1–1999:2.

Our second measure is obtained by using the method proposed by Ang et al. (2004). Hence we estimate a fourth-order VAR including $i_t^1, \pi_t, S_t, \Delta_4 y_t$ on the full sample, and we obtain ER_t^{APW} and TP_t^{APW} by recursively simulating forward the VAR for twenty periods for each data point from 1954:1 onwards. Following this method the two relevant measures are available for the sample 1955:1–2004:2. Finally, we use our own real-time measures, labelled ER_t^{RT} and TP_t^{RT}, derived using a fourth-order VAR.[7] As the first twenty years of observations are needed to estimate the first model and initialize the procedure, these measures are available only for the sample 1975:1–2004:2.

We report the three alternative decompositions of the yield spread in Figure 4, while their correlations are reported in Table 1. The three different measures of ER and the three different measures of TP are all positively correlated among them, although the correlation is not high. We take this as evidence of the presence of a common underlying factor, which is evident also graphically, but also as a signal

---

[6]Although we focus on predicting one-year ahead GDP growth, the results are robust to predicting GDP growth at longer horizons.

[7]The results from the real-time decomposition are robust with respect to the lag order of the VAR. This is however not the case for the APW decomposition.

that measurement matters. The correlation between all the ER and the TP factors is negative. The ER factors as measured by the Hamilton and Kim (2002) and Ang et al. (2004) methods are positively correlated with the yield spread while the correlation between the spread and the TP factors is low. The pattern is different for the measurement in real time in which case is the TP factor that features a higher correlation with the yield spread.

Figures 5–7 report the coefficients (with two-standard error confidence bands) from recursive estimation of equation (13), after an initialization sample of 20 years. Thus the predictive model for one-year ahead GDP growth includes a constant, the nominal short-term interest rate, inflation, an ER factor and a TP factor, where we consider, in turn, the three alternative decompositions of the yield spread. The models based on the HK and APW decompositions are estimated recursively considering all the possible sample splits after an initialization sample from 1955:1 to 1974:4, while the model estimated with the real-time measure needs another twenty years at the beginning of the sample to initialize the forecasting procedure and hence is estimated recursively with an initialization sample from 1975:1 to 1994:4. In the Hamilton-Kim model the ER component is only weakly significant, while in the two other models it is never significant. The TP component is always significant, and the coefficients on the nominal short rate and inflation are significant (except for inflation in the Hamilton-Kim decomposition), and consistent with the theoretical aggregate demand model with a short-term real interest rate. All decompositions give a picture in which an increase in the real short term rate implies a contraction in future output growth. This evidence is stronger in the case of the APW and real-time decompositions than in the HK decomposition. Importantly, there is no role for the expected path of future short rates in predicting future output, but instead the term premium plays an important role in predicting future output.[8]

Why does decomposing the yield spread into an expectations-related component and a term premium generate a different sign on inflation in our predictive model for output? To answer this question, Figure 8 shows U.S. inflation and the average expected future short-term interest rate based on our real-time measure of the expectations-related term: $ER_t^{RT} + i_t^1 = \frac{1}{20}\sum_{j=0}^{19} \mathrm{E}_t i_{t+j}^1$. Figure 8 shows a strong positive correlation between expected monetary policy and inflation (the correlation

---

[8]Considering the lags in the effects of monetary policy on GDP, it is perhaps not surprising that the expectations-related component has no significant predictive power for one-year ahead GDP growth. However, we obtain similar results when predicting GDP growth at longer horizons.

coefficient is 0.84), which implies a negative correlation between the ER component and the real short-term interest rate. Hence a predictive model for GDP growth based on the spread, the nominal short-term interest rate and inflation might deliver a "wrong" or non-significant sign on the real short rate as a by-product of the negative correlation between real short-term policy rates and the ER component of the spread. For the same reason, a predictive model which includes only the two components of the spread but not the short-term nominal interest rate and inflation might attach a stronger significance to the ER component. So the implementation of the decomposition of the spread has an impact on the significance and the sign of coefficients of other variables of the predictive models and generates an interpretable pattern for the coefficients determining the effect of monetary policy on future growth.

But what about the predictive power of the model? Figure 9 reports actual and predicted output growth, where forecasts are recursively generated using the APW decomposition of the spread. Note that the root-mean-squared error (RMSE) is the lowest ever reported (compare with Figure 2) and that the model shows no major failure in predicting turning points in growth. In particular it does much better than the models analyzed in Figure 2 in predicting the 1990–91 recession.

To sum up we consider a full-sample regression analysis. Table 2 contains the results of estimation of seven different predictive models. Model 1 is the standard predictive relation in which one-year ahead growth is related to the yield spread and a constant, Models 2 and 3 augment the specification of Model 1 by including progressively the nominal short-term rate and annual inflation, Model 4 is the specification similar to an aggregate demand model in which future output growth is related to the monetary policy stance as captured by the nominal short-term interest rate and inflation, and Model 5 features the same specification of Model 4 but the short-term interest rate is substituted with the long-term interest rate. Model 6 augments the basic aggregate demand specification with the spread decomposed into the ER and the RP components. There are three versions of Model 6 which consider the three different decomposition of the spread. Finally, Model 7 considers the case in which only the ER and the TP components of the spread are included in the model (using the APW decomposition).

Several comments are in order. First, the highest adjusted R^2 is reached by model 6(RT) which uses the real-time decomposition of the spread into the ER and the TP components. In this specification the ER component is not significant

while the TP component is positive and significant, while the nominal interest rate and inflation are also significant with a negative and positive sign respectively. The adjusted R^2 delivered by this specification is twice as high as that delivered by the spread only (Model 1) and almost twice as high as the one delivered by the aggregate demand-type specification (Model 4) in which only the nominal short-term rate and inflation are included. Moreover, the decomposition allows to estimate much more precisely the coefficient on inflation, which is never significant in the models that do not use the decomposition of the spread.

Second, the decomposition in real time dominates the other two decompositions in terms of predictive power and interpretability of the coefficients. Although the coefficients from the APW and the real-time decompositions are very similar, the adjusted R^2 from the latter is much higher. Hamilton and Kim (2002) found that both the ER and TP components are significant in predicting output growth, while Ang et al. (2004) found that only the ER term is significant. A comparative analysis of Models 6(APW) and 7 shows that the significance of the ER term tends to disappear when the nominal interest rate and inflation are included in the specification.

Finally, Figure 10 shows how decomposing the yield spread affects the recursive predictive power of the models, by adding Models 6(APW) and 6(RT) to the two best-performing models studied in Section 2. Using the APW decomposition considerably increases the predictive power compared with the best models without the decomposed yield spread. An additional increase in the predictive power is obtained by using the real-time decomposition, although estimates for this model are available only for the last part of the sample.

If we focus on our favorite decomposition, the one in real time, Table 2 shows that adding the term premium to the usual specification adopted in small macro models improves considerably the predictive performance of the model and also sharpens the precision with which the effects of the monetary policy stance on output is estimated. However, the positive and significant impact of the term premium on future growth in our preferred model needs some interpretation. We devote the next section to this issue.

4 A structural interpretation

The large literature that has examined the role of the yield spread in predicting future economic activity, and found it significant, has shown a clear tendency for interpreting the predictive power as dependent on the importance of the spread as a leading indicator for future monetary policy, which in turn is positively related to future economic activity. The two papers using an explicit decomposition of the yield spread into the ER and the TP components provide some evidence in favor of this interpretation. However, our results show that including the contemporaneous monetary policy stance makes the ER component lose its significance in predicting future GDP growth. Hence we are left with the problem of finding some motivation for the positive role of the term premium in predicting future economic activity.

Differently from Ang et al. (2004), Hamilton and Kim (2002) find a significant role for the TP component in predicting output growth. However, these authors encounter problems in finding a structural interpretation for their result. They propose a simple model based on the time variation in the variance of policy rates. According to the proposed two-factor affine model of the term structure (see, for example, Campbell et al. (1987)), an increase in interest rate volatility at the end of an expansion could explain why the term premium falls at the end of the expansion and therefore a low term premium predicts low future output growth. Hamilton and Kim (2002) find that volatility in monetary policy has some explanatory power for the term premium but, unfortunately, cyclical movements in volatility do not have the impact on the term premium predicted by theory, and therefore they are not able to account for the usefulness of the term premium for forecasting GDP growth.

Our interpretation of the positive impact of the term premium on GDP growth is somewhat different: we see our best forecasting model as a reduced form of an aggregate demand relation in which monetary policy has a delayed effect on output which is non-linear. Such non-linearity depends on the fact that the impact of monetary policy on output is a function of the time-varying risk aversion of agents: when risk aversion is high, and hence the term premium is large, monetary policy has less power in determining output fluctuations than when the risk aversion, and hence the term premium, is low.

We provide some evidence on our proposed interpretation by looking first at reduced-form estimation of a forecasting model and by then explicitly considering a structural interpretation. Our reduced-form evidence is illustrated in Table 3, using

13

the full sample 1975:1–2003:2. Table 3 reports in the first column the results of estimation of model 6(RT) and in the second column the results from estimation of our preferred predictive model for GDP growth:

$$\Delta_4 y_{t+4} = \alpha_0 + \alpha_2 TP_t^{RT} + \alpha_3 i_t^1 + \alpha_4 \pi_t + u_{1t}, \tag{14}$$

in which we have excluded from the specification the ER_t^{RT} factor, which was not significant. The third column in Table 3 contains the results of a re-specification of the prediction regression which illustrates the positive significance of the term premium in predicting future growth and the higher precision in measuring the effects of monetary policy generated by the inclusion of TP in the forecasting model:

$$\Delta_4 y_{t+4} = \alpha_0 + \left(\alpha_3 + \alpha_5 TP_t^{RT}\right) i_t^1 + \left(\alpha_4 + \alpha_6 TP_t^{RT}\right) \pi_t + u_{2t}. \tag{15}$$

From the second and third columns of Table 3 we note that the respecified model does not lead to any important reduction in the forecasting performance but it helps in interpreting the role of the term premium in predicting output. In fact, the significance of the term premium is explained in terms of the impact of monetary policy on output: monetary policy has a stronger impact on output growth when the term premium is low.

These results are interesting but they call for a structural interpretation. Columns 4 and 5 of Table 3 provide some evidence in this direction. Our structural model of reference is the intertemporal Euler equation for output given by (see, for example, Fuhrer and Rudebusch (2004))

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \sigma \left[i_t^1 - E_t \pi_{t+1} \right] + u_{3t}, \tag{16}$$

where $\tilde{y}_t$ is the output gap (the log deviation of output from potential, for which we use the measure from the Congressional Budget Office, available in the FRED database), $E_t \tilde{y}_{t+1}$ is the expectation formed at time t of the output gap at time $t+1$, i_t^1 is the nominal three-month interest rate, $E_t \pi_{t+1}$ is expected future inflation, and u_{3t} represents an aggregate demand shock. Naturally, the validity of this equation is limited to an economy without capital, durable goods investment, foreign trade and government spending, in which case output equals consumption and the output dynamics is determined by the Euler equation for the intertemporally optimal consumption choice. With the appropriate functional form for the underlying util-

14

ity function, the parameter σ can be interpreted as the intertemporal elasticity of substitution which is equal to the inverse of the relative coefficient of risk aversion.

As shown by a number of authors (see, for example, Estella and Fuhrer (2002)), simple descriptions of output dynamics as equation (16) are not very successful in matching the key dynamic features of the data and in pinning down the impact of monetary policy on output. This is confirmed in column 5 of Table 3, which reports GMM estimates of equation (16), where inflation is instrumented by lags of all the variables included in the model and the future output gap is projected from a VAR for the output gap, the spread, the short-term interest rate and inflation. The disappointing predictive performance of the model based on the monetary policy stance only can be then interpreted as the other side of the same coin. However, our evidence on output prediction suggests that a simple modification of the traditional structure, one that considers the possibility of time-varying risk aversion, may be more successful in making the Euler equation for output closer to the actual dynamics in the data.

To explore this possibility, we estimate the following specification:

$$\tilde{y}_t = \mathrm{E}_t \tilde{y}_{t+1} - \sigma_t \left[i_t^1 - \mathrm{E}_t \pi_{t+1} \right] + u_{4t}, \tag{17}$$

where

$$\sigma_t = \sigma_1 - \sigma_2 TP_t^{RT}, \tag{18}$$

so the intertemporal elasticity of substitution (and therefore the effect of monetary policy) is directly related to the term premium.[9] The results from this estimation, reported in the last column of Table 3, witness some success in this direction. The coefficients on the real interest rate are now significant and their sign are in line with the prediction of the theory: as agents become more risk averse, the effect of monetary policy on output becomes weaker.[10] Moreover, our best predictive model is easily interpreted as a reduced form of the forward-looking structure. In fact,

---

[9]Time-varying risk aversion is also key in using habit persistence to explain asset pricing puzzles, as in Campbell and Cochrane (1999). A similar idea to ours is proposed by Harvey (1988), although he uses the real yield spread to proxy for the time-varying risk aversion.

[10]As seen in Figure 4, all measures of the term premium show a strong decline from the early 1980s until the end of the sample. Our estimates of equation (17) then imply that monetary policy should have become more powerful during the last 20 years. This is consistent with the findings of Boivin and Giannoni (2003) that U.S. monetary policy has become more successful in stabilizing the economy after the early 1980s. (See also Cecchetti et al. (2004).)

equation (15) can be re-interpreted as a reduced form of equation (17), in which the rate of growth of potential output is proxied by a constant. The aggregate demand equation (17) provides an interesting framework for interpreting the forecasting performance of the different models that we have considered in the previous section.

5   Conclusions

In this paper we have tried to bring together two strands of the empirical macro literature: the reduced-form evidence that the yield spread helps in forecasting output and the structural evidence on the difficulties of estimating the effect of monetary policy on output in an intertemporal Euler equation. We have shown that the inclusion of a short-term interest rate and inflation improves the forecasting performance of the spread for future output but the coefficients on the short rate and inflation are difficult to interpret using a standard macroeconomic framework.

A decomposition of the yield spread into an expectations-related component and a term premium allows a better understanding of the forecasting model. In fact, the best forecasting model for output is obtained by considering the term premium, the short-term interest rate and inflation as a predictors. The expectations-related component loses its significance when it is considered jointly with the stance of monetary policy as a consequence of the high correlation between inflation and future expected monetary policy.

We provide a possible structural interpretation of these results by allowing for time-varying risk aversion, linearly related to our estimate of the term premium, in an intertemporal Euler equation for output. This simple modification of the standard aggregate demand framework allows us to pin down more precisely the impact of the policy stance on output in a forward-looking model for output fluctuations. Allowing for time-varying risk aversion is an avenue of research that is being currently explored by several strands in the international macro and international finance literature (see, for example, Dungey et al. (2000) or Kumar and Persaud (2002)).

Interestingly, the evidence that the impact of monetary policy on the business cycle is not constant over time but depends on other factors is in line with the verbal statements of monetary policymakers, although it has not yet been incorporated as

a non-linear effect in a structural macro model. In particular, there is an ongoing debate on the importance of fiscal discipline a pre-condition for successful inflation targeting (see, for example, Sims (2003)). In this debate the interaction between monetary and fiscal policy comes through the intertemporal budget constraint of the fiscal authority and its effect on expectations. Our specification of aggregate demand would allow for a more direct interaction between fiscal and monetary policy in the sense that fiscal fundamentals determine the term premium, which has an immediate effect on the impact of monetary policy on the business cycle. Of course, the non-linearity caused by the inclusion of the term premium in the output Euler equation makes the solution of a small macro model more complex and its effect can be evaluated only by simulation of an appropriately specified model. This is on our agenda for future research.

References

Ang, Andrew, Monika Piazzesi and Min Wei (2004), "What does the yield curve tell us about GDP growth?," mimeo, University of Chicago. Forthcoming, *Journal of Econometrics.*

Boivin, Jean and Marc P. Giannoni (2003), "Has monetary policy become more effective?" Working Paper No. 9459, National Bureau of Economic Research.

Campbell, John Y. and John H. Cochrane (1999), "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy* 107 (2), 205–251.

Campbell, John Y., Andrew W. Lo and A. Craig MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

Campbell, John Y. and Robert J. Shiller (1987), "Cointegration and tests of present value models," *Journal of Political Economy* 95 (5), 1062–1088.

Cecchetti, Stephen G., Alfonso Flores-Lagunes and Stefan Krause (2004), "Has monetary policy become more efficient? A cross country analysis," Working Paper No. 10973, National Bureau of Economic Research.

Dotsey, Michael (1998), "The predictive content of the interest rate term spread for future economic growth," *Federal Reserve Bank of Richmond Economic Quarterly* 84 (3), 31–51.

Dungey, Mardi, Vance L. Martin and Adrian R. Pagan (2000), "A multivariate latent factor decomposition of international bond yield spreads," *Journal of Applied Econometrics* 15 (6), 697–715.

Elton, Edwin J. (1999), "Expected return, realized return and asset pricing tests," *Journal of Finance* 54 (4), 1199–1220.

Estrella, Arturo (2004), "Why does the yield curve predict output and inflation?" mimeo, Federal Reserve Bank of New York.

Estrella, Arturo and Jeffrey C. Fuhrer (2002), "Dynamic inconsistencies: Counterfactual implications of a class of rational expectations models," *American Economic Review* 92 (4), 1013–1028.

Estrella, Arturo and Gikas A. Hardouvelis (1991), "The term structure as a predictor of real economic activity," *Journal of Finance* 46 (2), 555–576.

Estrella, Arturo, Anthony P. Rodrigues and Sebastian Schich (2003), "How stable is the predictive power of the yield curve? Evidence from Germany and the United States," *Review of Economics and Statistics* 85 (3), 629–644.

Feroli, Michael (2004), "Monetary policy and the information content of the yield spread," *Topics in Macroeconomics* 4 (1), Article 13.

Fuhrer, Jeffrey C. and Glenn D. Rudebusch (2004), "Estimating the Euler equation for output," *Journal of Monetary Economics* 51 (6), 1133–1153.

Goodhart, Charles and Boris Hofmann (2004), "Monetary policy transmission in simple backward looking models: The IS puzzle," mimeo, London School of Economics.

Hamilton, James D. and Dong Heon Kim (2002), "A reeximination of the predictability of economic activity using the yield spread," *Journal of Money, Credit, and Banking* 34 (2), 340–360.

Harvey, Campbell R. (1988), "The real term structure and consumption growth," *Journal of Financial Economics*, 22 (2), 305–333.

Harvey, Campbell R. (1989), "Forecasts of economic growth from the bond and stock markets," *Financial Analysts Journal*, 45, 38–45.

Kumar, Manmohan S. and Avinash Persaud (2001), "Pure contagion and investors' shifting risk appetite: Analytical issues and empirical evidence," *International Finance* 5 (3), 401–436.

Plosser, Charles I. and K. Geert Rouwenhorst (1994), "International term structures and real economic growth," *Journal of Monetary Economics* 33 (1), 133–155.

Sims, Christopher A. (2003), "Limits to inflation targeting," mimeo, Princeton University.

Stock, James H. and Mark W. Watson (1989), "New indexes of coincident and leading indicators," in Olivier Blanchard and Stanley Fischer (eds.), *NBER Macroeconomics Annual*, Cambridge, MA: The MIT Press.

Woodford, Michael (2003), *Interest and Prices*, Princeton, NJ: Princeton University Press.

Table 1: Correlation across alternative decompositions of the yield spread

	S_t^{20}	ER_t^{HK}	ER_t^{APW}	ER_t^{RT}	TP_t^{HK}	TP_t^{APW}	TP_t^{RT}
S_t^{20}	1	0.57	0.53	0.21	0.12	0.13	0.46
ER_t^{HK}		1	0.72	0.80	-0.39	-0.44	-0.39
ER_t^{APW}			1	0.63	-0.20	-0.74	-0.20
ER_t^{RT}				1	-0.77	-0.58	-0.77
TP_t^{HK}					1	0.61	0.79
TP_t^{APW}						1	0.62
TP_t^{RT}							1

Table 2: Comparison of alternative models for predicting GDP growth

	Model								
	1	2	3	4	5	6(HK)	6(APW)	6(RT)	7
Constant	2.00	3.30	3.30	4.77	4.44	4.28	5.68	5.21	2.21
	(0.28)	(0.60)	(0.60)	(0.38)	(0.57)	(0.64)	(0.92)	(0.82)	(0.50)
i_t^{20}	0.80	0.62	0.62		−0.11				
	(0.14)	(0.17)	(0.17)		(0.10)				
i_t^1	−0.80	−0.79	−0.80	−0.28		−0.41	−0.89	−0.78	
	(−)	(0.14)	(0.14)	(0.08)		(0.15)	(0.20)	(0.22)	
π_t			0.02	0.04	−0.10	0.17	0.43	0.41	
			(0.09)	(0.10)	(0.12)	(0.09)	(0.12)	(0.15)	
ER_t						0.56	−0.31	−0.42	0.83
						(0.17)	(0.24)	(0.33)	(0.18)
TP_t						0.70	0.68	0.48	0.68
						(0.20)	(0.28)	(0.16)	(0.28)
$\bar{R}^2$	0.182	0.239	0.22	0.126	0.045	0.31	0.233	0.354	0.186

Note: Heteroscedasticity-consistent standard errors within brackets. All estimates are based on the sample 1975:1–2003:2, with the exception of Model 6(HK) which has been estimated over the sample 1975:1–1999:2.

Table 3: Re-interpreting the best predicitive model for GDP growth

	Model 6(RT)	Eq. (14)	Eq. (15)	Eq. (16)	Eq. (17)
α_0	5.21	4.36	5.00		
	(0.82)	(0.37)	(0.37)		
α_1	-0.42				
	(0.33)				
α_2	0.48	0.66			
	(0.16)	(0.15)			
α_3	-0.78	-0.59	-0.67		
	(0.22)	(0.12)	(0.09)		
α_4	0.41	0.30	0.13		
	(0.15)	(0.10)	(0.03)		
α_5			0.36		
			(0.11)		
α_6			-0.12		
			(0.06)		
σ				-0.002	
				(0.02)	
σ_1					0.16
					(0.06)
σ_2					0.06
					(0.02)
$\bar{R}^2$	0.354	0.344	0.334		

Note: Heteroscedasticity-consistent standard errors within brackets. All estimates are based on the sample 1975:1–2003:2.

Model 6(RT) is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 ER_t^{RT} + \alpha_2 TP_t^{RT} + \alpha_3 i_t^1 + \alpha_4 \pi_t + u_t$

Equation (14) is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_2 TP_t^{RT} + \alpha_3 i_t^1 + \alpha_4 \pi_t + u_t$

Equation (15) is $\Delta_4 y_{t+4} = \alpha_0 + \left(\alpha_3 + \alpha_5 TP_t^{RT}\right) i_t^1 + \left(\alpha_4 + \alpha_6 TP_t^{RT}\right) \pi_t + u_t$

Equation (16) is $\tilde{y}_t = \mathrm{E}_t \tilde{y}_{t+1} - \sigma \left[i_t - \mathrm{E}_t \pi_{t+1}\right] + u_{3t}$

Equation (17) is $\tilde{y}_t = \mathrm{E}_t \tilde{y}_{t+1} - \left(\sigma_1 - \sigma_2 TP_t^{RT}\right) \left[i_t - E_t \pi_{t+1}\right] + u_{4t}$

Figure 1: Adjusted R^2 of alternative predictive models for GDP growth
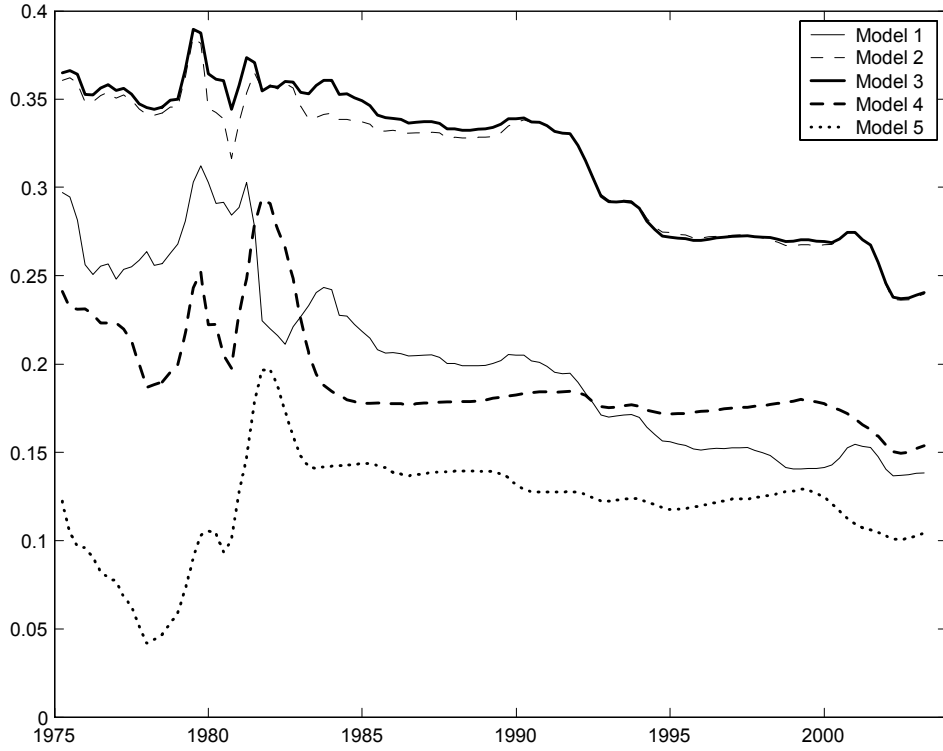
Model 1 is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 S_t^{20} + u_t$

Model 2 is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 S_t^{20} + \alpha_2 i_t^1 + u_t$

Model 3 is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 S_t^{20} + \alpha_2 i_t^1 + \alpha_3 \pi_t + u_t$

Model 4 is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 i_t^1 + \alpha_2 \pi_t + u_t$

Model 5 is $\Delta_4 y_{t+4} = \alpha_0 + \alpha_1 i_t^{20} + \alpha_2 \pi_t + u_t$

Figure 2: Actual and predicted GDP growth in alternative predictive models
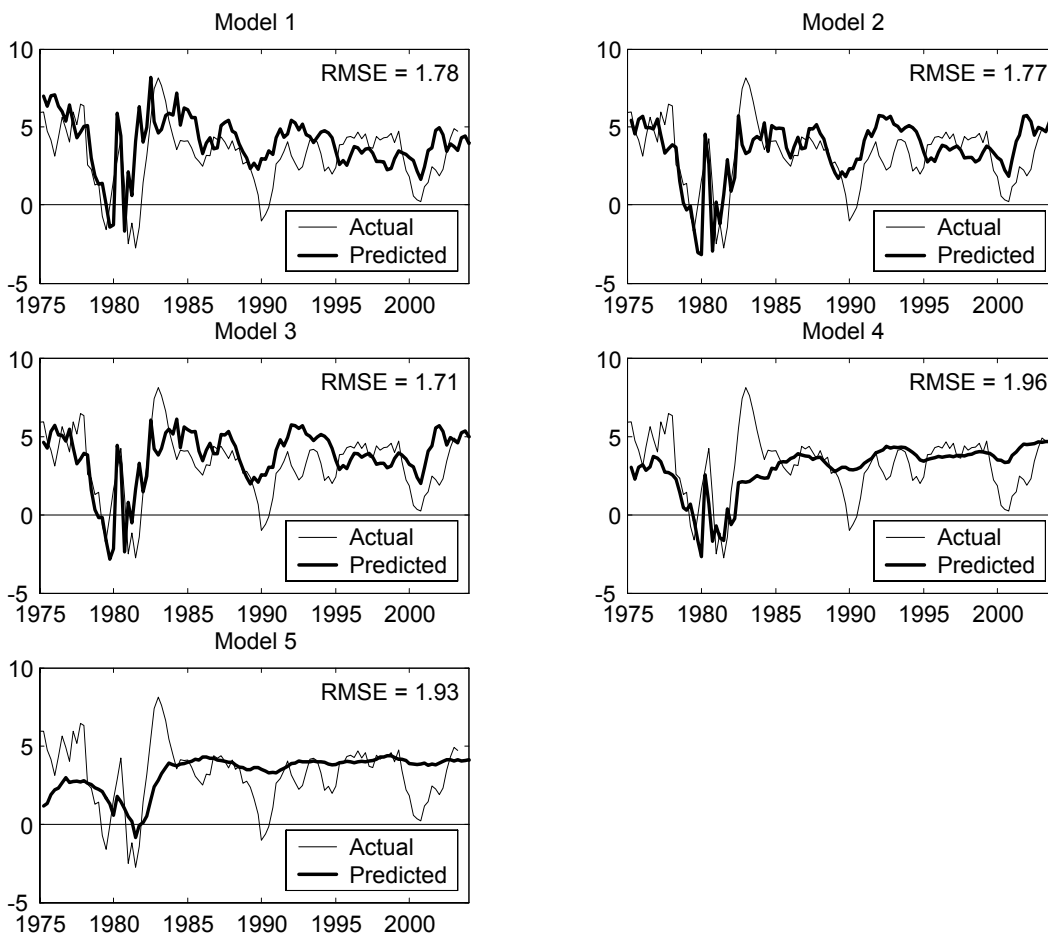
Figure 3: Recursive coefficient estimates in the best predictive model (Model 3) for GDP growth

25
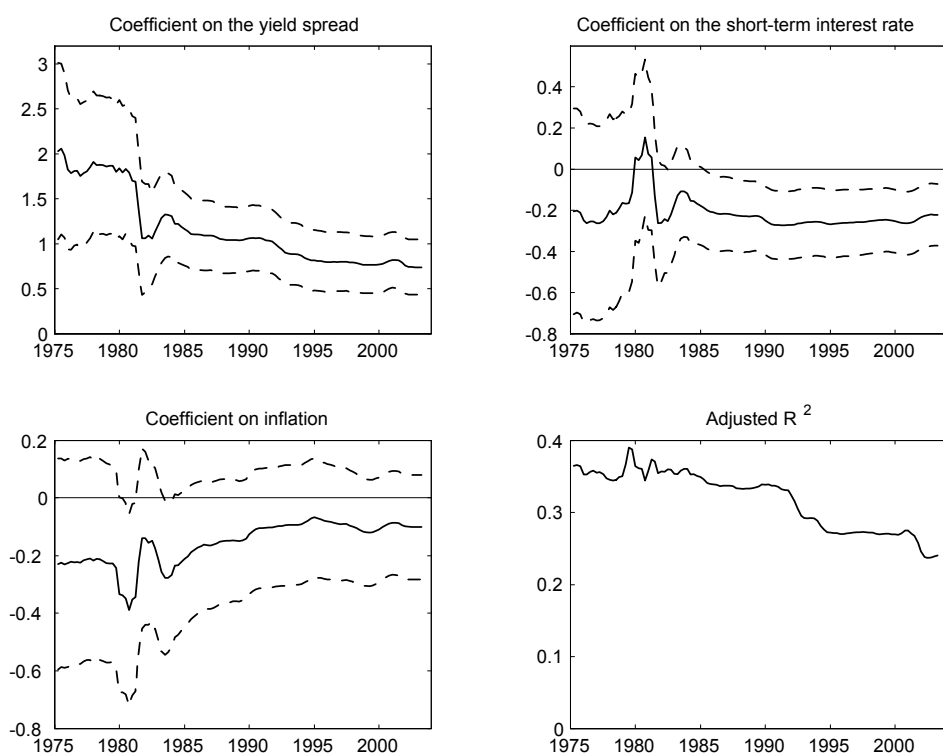
Figure 4: The yield spread and its three decompositions

The yield spread

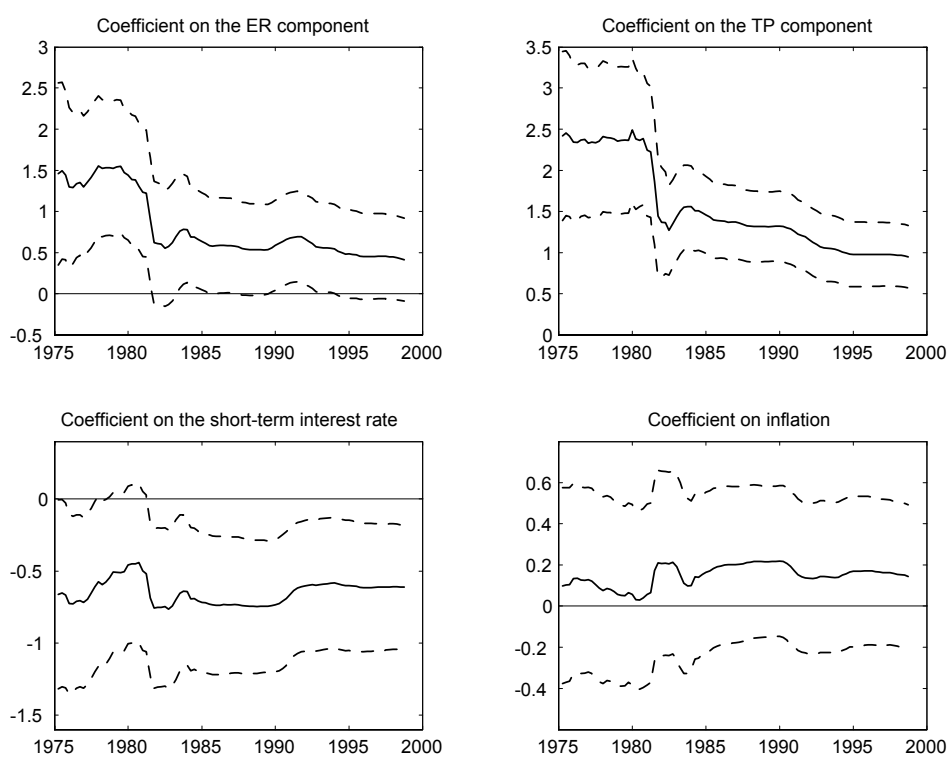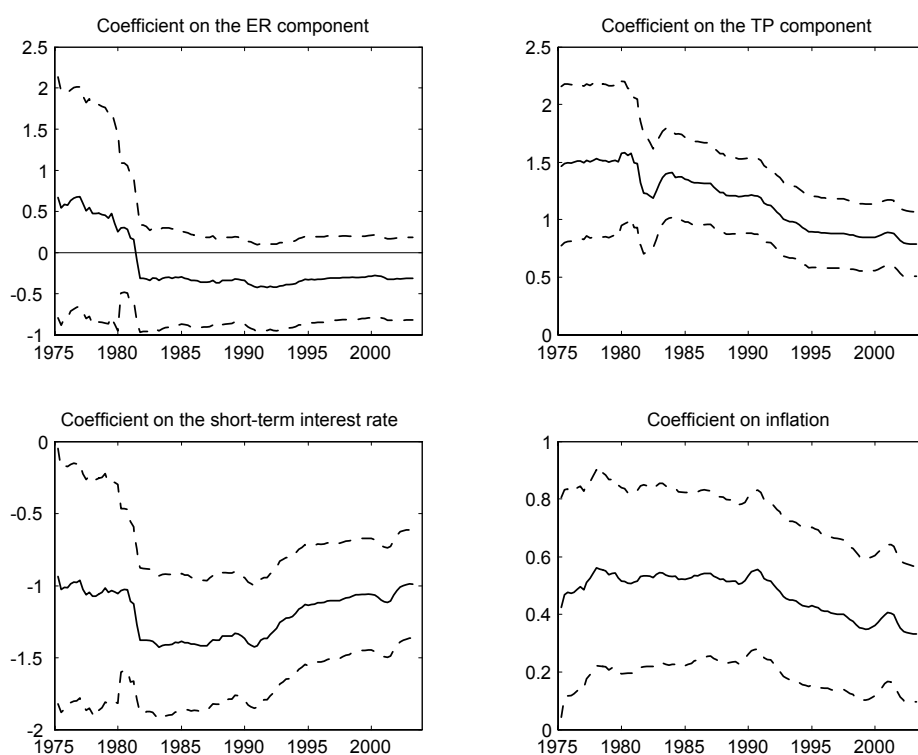The expectations-related component (ER)

HK
APW
RT

The term premium (TP)

HK
APW
RT
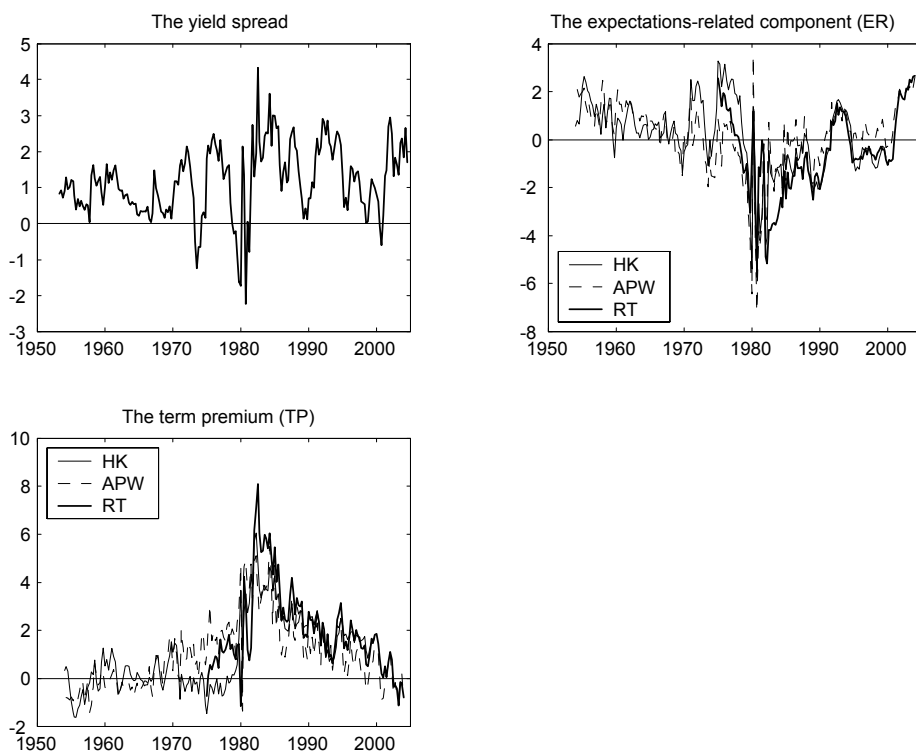
Figure 5: Recursive coefficient estimates in the predictive model for GDP growth based on the HK decomposition

Figure 6: Recursive coefficient estimates in the predictive model for GDP growth based on the APW decomposition

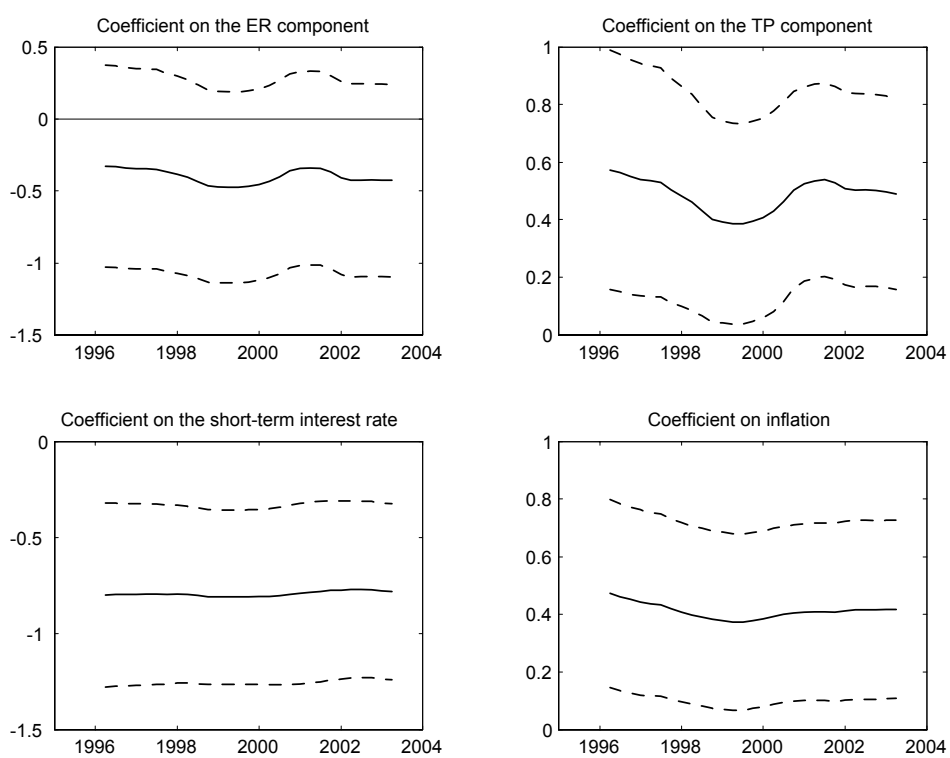Figure 7: Recursive coefficient estimates in the predictive model for GDP growth based on the real-time decomposition

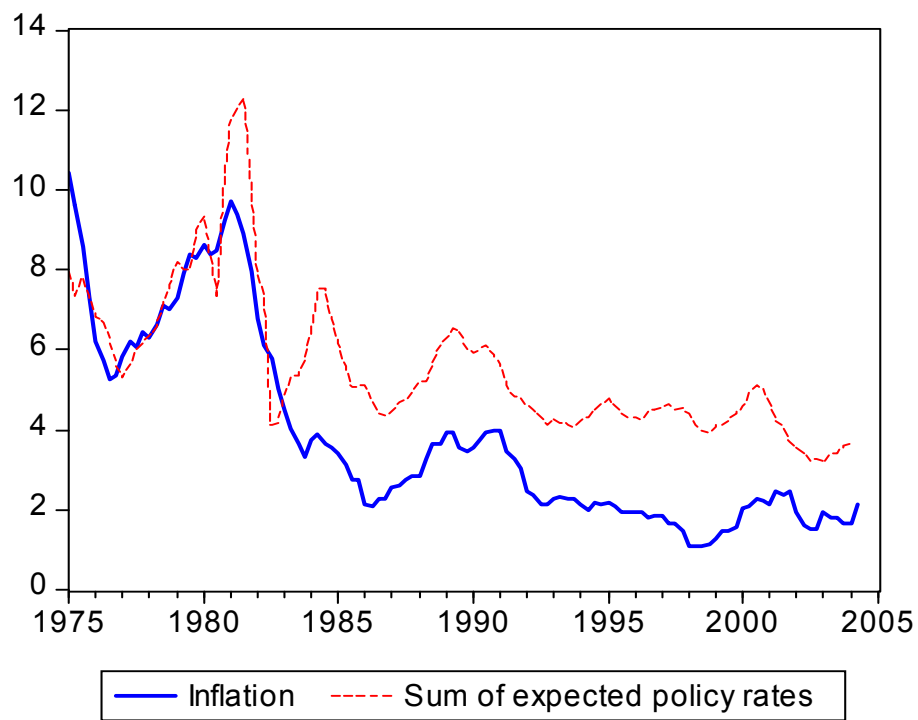Figure 8: Inflation and the sum of expected short-term interest rates $\left(ER_t^{RT} + i_t^1 \right)$

Figure 9: Actual and predicted GDP growth in a model including the short-term interest rate, inflation and the APW measure of the term premium
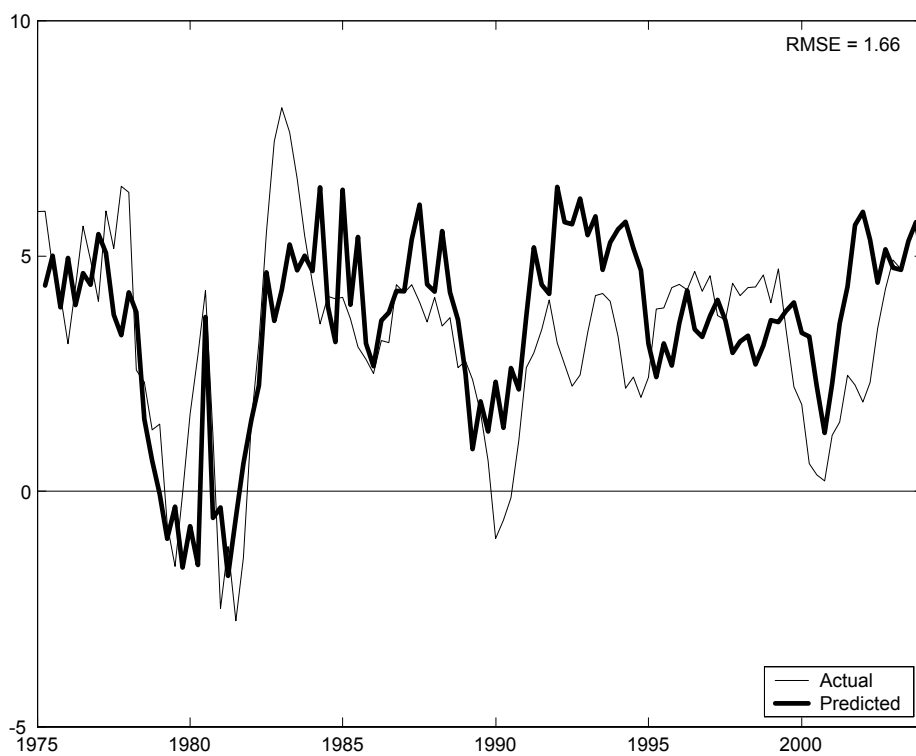

31

Figure 10: The effect on the predictive performance of including different measures of the term premium